

### Singuläre und multiple Imputation fehlender Einkommenswerte: ein empirischer Vergleich

Papastefanou, Georgios; Wiedenbeck, Michael

Veröffentlichungsversion / Published Version  
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:  
GESIS - Leibniz-Institut für Sozialwissenschaften

#### Empfohlene Zitierung / Suggested Citation:

Papastefanou, G., & Wiedenbeck, M. (1998). Singuläre und multiple Imputation fehlender Einkommenswerte: ein empirischer Vergleich. *ZUMA Nachrichten*, 22(43), 73-89. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-208245>

#### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

# SINGULÄRE UND MULTIPLE IMPUTATION FEHLENDER EINKOMMENSWERTE. EIN EMPIRISCHER VERGLEICH

GEORGIOS PAPASTEFANOU UND MICHAEL WIEDENBECK

Die Messung des Haushaltseinkommens ist besonders stark vom Problem des item-nonresponse betroffen. So beträgt der Anteil der fehlenden Werte im ALLBUS 1996 trotz Nachfrage mit Einkommenskategorien 26 (Westdeutschland) bzw. 19 Prozent (Ostdeutschland). Beim Eurobarometer 1992 liegt der Ausfall zwischen 6 und fast 50 Prozent. In der Forschungspraxis wird der Datenausfall oft entweder ignoriert oder durch einfache Imputationen, d.h. durch Zuweisung von Mittelwerten korrigiert. Diese Vorgehensweisen werden von Rubin (1987) in Frage gestellt, der statt dessen eine multiple Imputation befürwortet. Mit dieser Methode werden die fehlenden Werte mehrfach rekonstruiert, wobei jedesmal die Unsicherheit der auf einer Regressionsgleichung basierenden Einkommenschätzung berücksichtigt wird. Das vorliegende Papier untersucht am Beispiel der Eurobarometerdaten von 1992, welche Folgen eine einfache versus multiple Rekonstruktion der fehlenden Einkommenswerte für die Schätzung des Einkommenseffektes auf die Lebenszufriedenheit hat.

Income measurement is plagued by the problem of item nonresponse. For example, in the 1996 ALLBUS survey 26 percent of the sample in western Germany and 19 percent of the sample in eastern Germany have missing data on income. In Eurobarometer 1992 the percentage of missing data on income ranges between 6 and 50 percent across countries. Research using income as an explanatory variable usually ignores this problem or tries to solve it by replacing missing values by some form of means estimation. This procedure, called simple imputation, differs from the multiple imputation proposed by Rubin (1987). In multiple imputation, holes in the data are also filled by regression estimated means. The procedure is repeated several times, taking into account the regression error each time. Using Eurobarometer 1992 data, the paper compares the consequences of using simple imputation versus multiple imputations to estimate the effect of income on life satisfaction ratings.

## 1. Einleitung

Trotz fortgeschrittener Erhebungsverfahren ist in der empirischen Sozialforschung damit zu rechnen, daß bei bestimmten Erhebungsmerkmalen Werte zu einem erheblichen Anteil fehlen. Von diesem als item-nonresponse bekannten Problem (Little/Rubin 1987) ist die Messung des Einkommens in besonderer Weise betroffen. So weisen die demografischen Daten des ALLBUS von 1994 zwar Anteile von fehlenden Werten auf, die überwiegend unter einem Prozent liegen. Die Einkommensvariablen (Haushaltseinkommen und persönliches Monatsnettoeinkommen) haben jedoch mit 14 Prozent den größten Anteil fehlender Angaben. Betrachtet man die Einkommensmessung im ALLBUS seit seinem Beginn 1980, dann ist festzustellen, daß der Anteil fehlender Werte in Westdeutschland auch bei Nachfrage mit Einkommenskategorien zwischen 12 Prozent (1986) und 40 Prozent (1982) schwankt. Ein eindeutiger Trend ist dabei nicht auszumachen, nach relativ niedrigen Werten (15- bis 18 Prozent) in den neunziger Jahren liegt der Anteil fehlender Einkommenswerte beim ALLBUS 1996 bei 26 Prozent. In Ostdeutschland hat sich der Ausfall bei der Einkommensmessung von 10- bis 12 Prozent in den ersten drei Erhebungen 1991, 1992 und 1994 auf 19 Prozent im Jahr 1996 erhöht.

Ansonsten ist bei demografischen Variablen nur noch die Messung der beruflichen und schulischen Ausbildung von Vater und Mutter von einem nennenswerten Datenausfall betroffen, der zwischen drei und sieben Prozent liegt. Insofern ist es bezeichnend, daß in statistischen Abhandlungen die Problematik fehlender Werte meist am Beispiel der Einkommensmessung erörtert wird (z. B. Rubin 1987; Little/Schenker 1995).

Zur Schätzung des Einkommenseffektes dienen meist Daten aus nationalen Repräsentativumfragen, in denen neben anderen Informationen das Haushaltseinkommen und die Lebenszufriedenheit erfaßt werden. Beim Eurobarometer kann man beispielsweise feststellen, daß der Anteil fehlender Einkommensangaben durchschnittlich 22 Prozent beträgt und zwischen sechs Prozent (bei portugiesischen Daten) und bis zu 50 Prozent (bei nordirischen Daten) variiert. Bei anderen öffentlich zugänglichen Daten aus regelmäßig durchgeführten Erhebungen, wie z.B. ISSP, oder Wohlfahrtssurvey sind ebenfalls hohe nonresponse-Raten der Einkommensvariable festzustellen.

Empirische Forschungsarbeiten, die das Einkommen als erklärende Variable verwenden, scheinen in einem hohen Anteil fehlender Werte oft kein Problem zu sehen. In der World Happiness Database (Veenhoven 1994) z.B. sind über 140 Studien zum Einkommen als Determinante der Lebenszufriedenheit dokumentiert, aber es werden keine Angaben über den Ausfall in der Einkommensmessung berichtet. In einer Durchsicht von Einkommensanalysen mit ALLBUS-Daten, war ebenfalls ein nur geringes Problembewußtsein festzustellen (Papastefanou/Fleck 1995). Von den 20 ausgewählten Analysen mit Ein-

kommensvariable äußern die meisten nichts zu dieser Problematik. Zwei Veröffentlichungen, die das Problem fehlender Werte ansprechen, lösen es durch Beschränkung der Analyse auf Fälle mit gültigen Einkommensangaben.

In dem Maße wie Standard-Statistiksoftwarepakete eine einfache Schließung der Datenlücken mit Ersatzwerten (Imputation) erlauben, bietet sich eine scheinbar einfache Problemlösung. Allerdings gibt es neuere Arbeiten, die nachdrücklich auf die Problematik verschiedener Varianten der Imputation hinweisen (Reinecke/Schmidt 1994). Wir schließen uns dieser Zielsetzung mit der folgenden Analyse an, und versuchen am Beispiel von Daten des Eurobarometer 1992 aufzuzeigen, mit welchen Konsequenzen man bei der empirischen Analyse zu rechnen hat, je nachdem ob man eine singuläre und eine multiple Imputation (wie sie z. B. Rubin (1987) propagiert) vornimmt. Hierzu generieren wir die fehlenden Einkommenswerte mit Hilfe einer singulären und einer multiplen Imputationsstrategie. Auf der Basis eines Datensatzes, der beobachtete und imputierte (singulär versus multiple) Einkommensdaten umfaßt, schätzen wir den Effekt des Haushaltseinkommens auf die Lebenszufriedenheit und zeigen, daß sich je nach Imputationsverfahren qualitativ andere inhaltliche Interpretationen ergeben können.

## 2. Vorgehensweise bei singulärer und multipler Imputation

Zur Bestimmung des zu imputierenden (Ersatz)-Wertes bei der sogenannten singulären Imputation gibt es verschiedene Verfahren, um die im Datensatz verfügbaren Informationen zu nutzen. Ein einfaches Verfahren besteht z. B. darin, den Personen mit fehlendem Einkommen, das mittlere Einkommen der Personengruppe mit vollständigen Angaben zuzuweisen. Dieses Verfahren wird singuläre Imputation genannt und ist mittlerweile in Standardstatistikprogrammpaketen wie SPSS, STATISTICA, EQS als Menüpunkt verfügbar.

Eine Verfeinerung dieser Verfahrensweise besteht darin, bedingte Mittelwerte zu imputieren. Nach bestimmten Merkmalen abzugrenzende Personengruppen mit fehlender Einkommensangabe erhalten dabei das mittlere Einkommen jener nach den gleichen Merkmalen gebildeten Personengruppen mit vollständigen Einkommenswerten. Eine weitere Verfeinerung der singulären Imputation besteht darin, die fehlenden Werte durch Prognosewerte zu ersetzen, die mit Hilfe geeigneter Regressionsgleichungen errechnet worden sind (Yates 1963).

Vom Verfahren der singulären Imputation grenzt sich die sogenannte multiple imputation (Rubin 1987, 1995) ab. Der Ansatz der multiplen Imputation geht davon aus, daß ein singulär imputierter Wert nicht die Unsicherheit repräsentieren kann, die bei der Schät-

zung von Ersatzwerten durch z. B. eine Einkommensgleichung gegeben ist (Rubin/Schenker 1991: 586). Eine Datenmatrix, die aus einfach imputierten Einkommenswerten und beobachteten Werten besteht, weist eine methodisch reduzierte Varianz der Einkommensvariable auf, die die tatsächliche Variabilität um so schlechter widerspiegelt, je höher der Anteil der fehlenden Werte ist.

Eine multiple Imputation ist deshalb darauf gerichtet, die mit den imputierten Werten verbundene Unsicherheit bzw. Varianz, die sich letztendlich auf Effektgröße und Konfidenzintervall niederschlagen muß, bei der Parameterschätzung zu berücksichtigen. Die Strategie der multiplen Imputation umfaßt zwei aufeinander aufbauende Verfahrensschritte: im ersten Schritt wird die lückenhafte Datenmatrix auf der Basis einer geschätzten bedingten Verteilung gegeben die beobachteten Daten mehrfach komplettiert. Dadurch entstehen mehrere vollständige Datenmatrizen, die jede für sich einer Analyse unterzogen werden. Aus den verschiedenen Analysen ergibt sich eine Stichprobe von geschätzten Parametern, die anschließend zu einer einzigen Parameterschätzung zusammengefaßt werden. Die Variabilität der Parameter in dieser Stichprobe geht in die Berechnung der Standardfehler ein und repräsentiert die zusätzliche Unsicherheit aufgrund der fehlenden Werte.

Durch multiple Imputation werden also Datenausfälle nicht durch einzelne Werte ersetzt, die in Auswertungen wie tatsächlich erhobene Daten behandelt werden, sondern durch Verteilungen, die durch Stichproben repräsentiert werden. Ob durch diese Substitutionen eine Datenbasis gewonnen wird, die valide Ergebnisse ermöglicht, hängt u.a. von den statistischen Eigenschaften des Ausfallprozesses ab. Eine wichtige Annahme zur Vereinfachung bei likelihood-basierten Analysen ist die Unabhängigkeit des Ausfallprozesses von den Werten der nicht-beobachteten Variablen sowie den Parametern der Verteilungsklasse, die in den Analysen verwendet werden. Die Ausfälle werden dann als missing at random (MAR) bezeichnet (Little/Rubin 1987: 89). Unter dieser Voraussetzung ist die Likelihood-Funktion der beobachteten Daten nicht vom Ausfallprozeß bzw. dessen Parametern abhängig. Der Ausfallprozeß kann dann bei den Analysen ignoriert werden. Die Ergänzung der beobachteten Daten zu quasi-vollständigen Datensätzen dient in erster Linie der technischen Vereinfachung der Analysen. Die imputierten Werte werden als Ziehung aus einer Verteilung gewonnen, wobei nach Möglichkeit auf beobachtete Einflußgrößen bedingt wird.

Im Rahmen der hier betrachteten inhaltlichen Fragestellung bedeutet dies, daß das beobachtete Haushaltseinkommen als lineare Funktion der Faktoren Geschlecht, höherwertige Haushaltsausstattung, berufliche Position, Bildungsniveau und Altersgruppe gesehen wird, und zwar für jedes Land getrennt. Danach werden für die Personen, die keine gülti-

gen Einkommensangaben aufweisen, auf der Grundlage ihrer Ausprägungen bei den Prädiktoren und der geschätzten Regressionsparameter errechnet.

Wie auch bei der singulären Imputation wird bei der multiplen Imputation ein Regressionsmodell für die fehlenden Werte verwendet. Bei der singulären Imputation werden bedingte Mittelwerte entsprechend der geschätzten Regressionsgleichung als Ersatzwerte imputiert. Bei der multiplen Imputation sind dagegen die Ersatzwerte Zufallsziehungen aus den bedingten Verteilungen selbst.

Nach Rubin/Schenker (1991: 586) sind bereits drei Wiederholungen ( $m$ ) ausreichend („in typical problems choosing  $M = 3$  works very well“). In der vorliegenden Analyse haben wir die Imputation viermal repliziert und ein Regressionsmodell mit normalverteilten Fehlern verwendet. Die Parameterschätzungen der einzelnen Replikationen werden schließlich zusammengefaßt. Wir folgen hierbei der Prozedur von Rubin.

### **3. Schätzung des Effektes von Haushaltseinkommen auf die Lebenszufriedenheit mit ergänzten Einkommensdaten**

#### **3.1 Substantielle Fragestellung und Analysemodell**

In der Forschungsdiskussion über den Einfluß des Haushaltseinkommens auf die allgemeine Lebenszufriedenheit stehen sich zwei z. T. einander widersprechende Hypothesen gegenüber.

Inglehart (Inglehart/Rabier 1986; Inglehart 1989) hat die Hypothese entwickelt, daß sozio-ökonomische Strukturen tendenziell *keinen* länger andauernden Effekt auf die Lebenszufriedenheit haben. Er geht davon aus, daß Personen in sozio-ökonomischen Lagen ihre Ansprüche bzw. Erwartungen an den Ist-Zustand anpassen. Personen in deprivierten Lagen sind anfangs zwar unzufrieden, mit der Zeit sinken ihre Ansprüche und ihre Unzufriedenheit wird geringer. Bei Personen, die sich längere Zeit in einer guten Lage befinden, steigen hingegen die Ansprüche, und ihre Zufriedenheit sinkt. Aufgrund dieser mit wachsender Zustandsdauer erfolgenden Regression zur Mitte seien letztlich keine bzw. nur geringe sozio-ökonomischen Differenzierungen der Lebenszufriedenheit zu beobachten. Nur dann, wenn die Verweildauer in Deprivations- bzw. Privationslagen relativ kurz ist, wird hohe Unzufriedenheit bzw. Zufriedenheit gespürt und geäußert. Dementsprechend kann man nach Inglehart erwarten, daß das Haushaltseinkommen als relativ stabiler Indikator der materiellen Lage keinen differenzierenden Effekt auf die Lebenszufriedenheit hat.

Im Gegensatz dazu geht Veenhoven (1990, 1993, 1994) davon aus, daß dem materiellen Komfort ein großer Stellenwert für die Lebenszufriedenheit zukommt. Nach seiner Auf-

fassung schlägt sich neben kognitiven Bewertungen von Ist-Soll-Vergleichen eine hedonistisch-affektive Bewertungskomponente in der erfahrenen Lebenszufriedenheit nieder. Sie äußert sich bei Personen als ein *Gefühl* der Lebenszufriedenheit und als eher intuitives Bewerten oder Einschätzen seines Lebens. Wenn ein bestimmtes, notwendiges Maß von Lebensbedürfnissen befriedigt wird, erlebt das Individuum ein positives Lebensgefühl. Je besser grundlegende Lebensbedürfnisse durch gesellschaftliche Leistungen und Anforderungen (wie materieller Komfort, soziale Gleichheit, politische Freiheit und Zugang zu Wissen) befriedigt werden, um so zufriedener fühlen sich die Gesellschaftsmitglieder.

Das intuitive Wohlgefühl ist in der Sicht von Veenhoven (1993) eine Art evolutionär entwickeltes Frühwarnsystem bezüglich grundlegender Bedürfnisse nach Nahrung, Unterkunft, Sicherheit, Identität und Sinn. Materieller Komfort erlaubt die Befriedigung eines Teils solcher grundlegenden Bedürfnisse und wird sich deshalb in einem intuitiven Gefühl des Wohlergehens niederschlagen, das sich wiederum in einer allgemeinen Lebenszufriedenheit ausdrückt (Veenhoven 1994). Da in westlichen Konsumgesellschaften der materielle Komfort wesentlich vom Haushaltseinkommen abhängt, kann man diesbezüglich einen direkten Effekt auf die Lebenszufriedenheit erwarten.

Zur Überprüfung dieser einfachen Hypothesen ziehen wir Daten des Eurobarometer von 1992 heran, der Daten zur Lebenszufriedenheit sowie Informationen über die ökonomische und soziale Lage von Personen in 15 europäischen Ländern umfaßt (Kommission der Europäischen Gemeinschaften 1992). Im Eurobarometer wird die Lebenszufriedenheit mit einer Vierer-Skala erfaßt, für unsere Analyse haben wir die Antwortkategorien in zwei Werte zusammengefaßt und zwar „sehr zufrieden“ und „zufrieden“ auf der einen, und „unzufrieden“ sowie „sehr unzufrieden“ auf der anderen Seite. Wir vergleichen also überwiegend zufriedene ( $Z=1$ ) mit überwiegend unzufriedenen Personen ( $Z=0$ ).

Um mit dem Haushaltseinkommen adäquat die materielle Wohlfahrt von Personen zu indizieren, verwenden wir das bedarfsgewichtete Pro-Kopf-Einkommen, wie es von der OECD verwendet wird (Haagenaars/de Vos/Zaidi 1995). Es handelt sich bei dieser Meßgröße um die pro-Kopf-bezogene Einkommensposition des Haushaltes relativ zur Armutsschwelle, die wir mit 50 Prozent des durchschnittlichen Haushaltseinkommens angesetzt haben. Um mögliche nicht-lineare Effekte der materiellen Wohlfahrt auf die Lebenszufriedenheit zu messen, gruppieren wir die relativen Einkommenspositionswerte zu Intervallen. Wir gelangen so zur Unterscheidung von vier Personenkategorien mit unterschiedlichem materiellen Niveau, je nach Abstand des Haushaltes zur Armutsschwelle:

1. Personen in Haushalten mit Armut (unterhalb der Armutsschwelle) ( $X_0$ ),
2. mit bescheidenem (bis 50 Prozent oberhalb der Armutsschwelle - BEScheiden) ( $X_1$ ),
3. mit mittlerem (zwischen 51 und 100 Prozent oberhalb der Armutsschwelle - WOHLHABEND) ( $X_2$ ),
4. und großem materiellen Niveau (mehr als 100 Prozent oberhalb der Armutsschwelle - REICH) ( $X_3$ ).

Statistisch überprüfen wir die Einkommenshypothesen mittels der Schätzung eines logistischen Regressionsmodells,

$$\log(p(Z=1)/p(Z=0)) = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + e^1)$$

in dem  $p(Z)$  die Wahrscheinlichkeit, überwiegend zufrieden zu sein als abhängige Variable und  $X_1, X_2, X_3$  als erklärende Variablen den oben benannten Haushaltstyp darstellen. Als Referenzgruppe verwenden wir dabei die Personen in armen Haushalten.<sup>2)</sup>

### 3.2 Ein Einkommensmodell als Imputationsmodell

Beim grundlegenden Einkommensmodell<sup>3)</sup> zur Imputation fehlender Werte verwenden wir als Zielgröße die relative Einkommensposition des Haushalts und folgende Variablen als Prädiktoren: FERIENWO: 1=Besitz von Ferienwohnung, 0=Sonstige; GENDER: 1=männlich, 2=weiblich; NACH18: 1=Bildungsabschluß nach dem 18. Lebensjahr, 0=Sonstige; TRONIC: 1=Besitz von hochwertigen elektronischen Haushaltsgütern, 0=Sonstige; MANAGER: 1=Managerposition, 0=Sonstige; UNEMPLD: 1=arbeitslos, 0=Sonstige; A20BIS29: Altersgruppe 20-29 Jahre; A30BIS39: Altersgruppe 30-39 Jahre; A40BIS52: Altersgruppe 40-52 Jahre; A53BIS60: Altersgruppe 53-60 Jahre; A61BIS70: Altersgruppe 61-70 Jahre; A71UND: Altersgruppe 71 Jahre und älter, Referenzgruppe bei der Altersvariable sind die 18- und 19jährigen.

Dieses Einkommensmodell haben wir für jede einzelne Stichprobe der Eurobarometererhebung von 1992 getrennt geschätzt. Dabei ergaben sich in den einzelnen nationalen Stichproben ganz erheblich unterschiedliche Prognosegüten, wie sie z.B. an der Standardabweichung (STD) der Residuen (siehe Tabelle 2) sichtbar werden.

Die Standardabweichung bei der Voraussage der relativen Einkommensposition der Haushalte liegt zwischen 28 bis 52 Punkten bei einem Wertebereich der relativen Einkommensposition, der von minus 60 bis 250 reicht. Innerhalb der gesamten Spannweite des Standardfehlers ist der Standardfehler bei den Daten Ostdeutschlands und Luxemburgs eher klein. Bei den Datensätzen von Westdeutschland, Portugal, Norwegen, Spanien, Frankreich und Dänemark bewegt sich der Standardfehler zwischen 36 und 38 Punkten in einem mittleren Bereich. Griechenland, Niederlande und Italien gehören zu



den Ländern mit einem Standardfehler von über 40 bis 43. Die größte Unsicherheit bei der Voraussage der relativen Einkommensposition der Haushalte ist bei den Datensätzen der Länder Irland, Belgien, Nord-Irland und Großbritannien gegeben. Sie reicht von 46 bis 52 Punkten.

**Tabelle 2: Anteil fehlender Einkommenswerte (in Prozent ), sowie Variabilität der Einkommensresiduen in nationalen Stichproben des Eurobarometer 1992**

Land	Anteil fehlender Einkommenswerte in %	Variabilität der Einkommensresiduen (STD)	Land	Anteil fehlender Einkommenswerte in %	Variabilität der Einkommensresiduen (STD)
Portugal	6,5	35,98	Italien	28,9	43,53
Norwegen	8,7	36,06	Großbritannien	30,4	51,69
Ostdeutschland	8,8	27,26	Spanien	32,1	36,79
Dänemark	9,7	38,15	Belgien	36,8	48,12
Westdeutschland	10,8	34,58	Luxemburg	38,9	29,22
Niederlande	14,3	43,44	Irland	40,9	45,74
Frankreich	18,5	37,49	Nordirland	49,5	51,22
Griechenland	24,7	40,983			
Quelle: Eurobarometer 1992, eigene Berechnung					

### 3.3 Das Ausmaß an Imputationsunsicherheit in den einzelnen Stichproben

Die mit der Prognosegüte verbundene Unsicherheit des Imputationsmodells wird verstärkt durch den Anteil der fehlenden Werte. Der Anteil fehlender Einkommenswerte liegt beim Eurobarometer 1992 zwischen sechs und 50 Prozent (siehe Tabelle 3). In den Ländern Portugal, Norwegen, Ostdeutschland, Dänemark und Westdeutschland liegen eher moderate Ausfallraten vor zwischen sechs und elf Prozent. Bei einer zweiten Gruppe von Ländern liegt die Ausfallrate zwischen 14 und 18 Prozent (Niederlande und Frankreich). In den Daten von Griechenland, Italien, Großbritannien, und Spanien sind die Ausfälle zwischen 24 und 32 schon ziemlich beträchtlich. Sehr große Ausfälle (zwischen 36 und 50 Prozent) bezüglich der Haushaltseinkommensangaben weisen die Datensätze von Belgien, Luxemburg, Irland und Nord-Irland auf.

Kombiniert man nun Ausfallrate und Prognosestandardfehler als Quellen der Unsicherheit, so gelangt man zu einer Einordnung der nationalen Datensätze bezüglich ihrer Sen-

sibilität gegenüber der ausfallbedingten Unsicherheit bei der Schätzung von Einkommenseffekten. Wir unterscheiden zwischen hohem und geringem Ausfall sowie zwischen hohem und geringen Prognosefehler und gelangen zu folgender Kategorisierung, nach der insbesondere in den Datensätzen von Irland, Belgien, Nordirland und Großbritannien das Problem der Unsicherheit verschärft gegeben ist (siehe Tabelle 3).

**Tabelle 3: Einordnung der EG- Länder-Datensätze nach Ausmaß des Prognosefehler der Einkommensfunktion und Ausfallrate**

	Länder mit geringerem Prognosefehler (28 - 38)	Länder mit höherem Prognosefehler (40 - 52)
Länder mit geringerer Ausfallrate (9 - 10%)	Ostdeutschland, Westdeutschland, Portugal, Norwegen, Dänemark	
Länder mit mittlerer Ausfallrate (11 - 31%)	Frankreich	Niederlande, Griechenland Italien
Länder mit hoher Ausfallrate (32 - 49%)	Luxemburg, Spanien	Belgien, Irland, Nordirland, Großbritannien

### 3.4 Die Schätzung von Einkommenseffekten mit ergänzten Einkommensdaten

Es stellt sich nun die Frage, wie sich die Art des Imputationsverfahrens auf die Schätzung des Einkommenseffektes auf die Lebenszufriedenheit auswirkt. Zu welcher Schlußfolgerung hinsichtlich der Zufriedenheitsunterschiede zwischen armen, bescheidenen und reichen Haushalten gelangt man mit multipel imputierten im Vergleich zu singulär imputierten Daten? Betrachten wir die Ergebnisse zunächst für die Datensätze von Belgien, Irland, Nordirland und Großbritannien, jenen Ländern, die aufgrund des hohen Ausfallanteils und des hohen Standardfehlers der Einkommensprognose von der Imputationsunsicherheit am stärksten betroffen sind (siehe Tabelle 4).

Im Falle Großbritanniens würde man bei singulärer Imputation zum Ergebnis kommen, daß mit zunehmendem Wohlstand der Gewinn an Lebenszufriedenheit steigt. Personen in Haushalten, die über der Armutsschwelle liegen, sind zufriedener als Personen in armen Haushalten; Personen in reichen Haushalten sogar um das zehnfache gegenüber Personen in armen Haushalten. Aber auch Personen in Haushalten, die nur etwas über der Armutsschwelle liegen, erleben dies als einen Zufriedenheitsgewinn. Man könnte dieses Ergebnis als Hinweis sehen, daß in England die Armen in hohem Maße eine deprivierte Randlage erleben. Bei multipler Imputation erhalten wir jedoch Ergebnisse, die eine andere Interpretation nahelegen: Auf dieser Basis stellt man fest, daß Personen in armen und bescheidenen Haushalten sich hinsichtlich ihrer Lebenszufriedenheit nicht unterscheiden. Der

Zugewinn an Lebenszufriedenheit bei Personen in wohlhabenden gegenüber Personen in armen Haushalten ist nicht sehr groß, der Abstand zwischen Armen und Reichen ist zwar groß, aber nicht in dem Ausmaß wie mit singularär imputierten Daten. Diese Ergebnisse könnte man als Modifikation von Veenhovens These interpretieren: nicht die Deprivation existentieller Bedürfnisse bewirkt nachhaltige Unzufriedenheit, sondern ein großer Konsumspielraum drückt sich in entsprechend hoher Lebenszufriedenheit aus.

Bei den irischen Daten erhält man eine Inkonsistenz zwischen singularer und multipler Imputation, die gravierende Konsequenzen für die Interpretation des Brutto-Einkommenseffektes auf die Lebenszufriedenheit hat. Auf der Grundlage von singular imputierten Daten würde man schließen, daß sich Personen in armen und bescheidenen Haushalten hinsichtlich ihrer Lebenszufriedenheit nicht unterscheiden, erst mit zunehmendem Wohlstand würde sich die Lebenszufriedenheit erhöhen. Bei multipler Imputation jedoch sind die Zufriedenheitsunterschiede zwischen Personen in armen und bescheidenen, wohlhabenden und reichen Haushalten nicht signifikant. Die multiple Imputation würde die Inglehartsche These stützen, während singular imputierte Daten eher zu Ergebnissen führen, die für eine Variante der Veenhovenschen These sprechen.

Auch die Ergebnisse für Nord-Irland führen je nach Imputationsverfahren zu unterschiedlichen Schlußfolgerungen. Bei singularer Imputation erhält man den Befund, daß Personen in allen Haushaltsgruppen über der Armutsschwelle signifikant höhere Zufriedenheitsanteile aufweisen als Personen in armen Haushalten, wobei die wohlhabenden den größten relativen Zugewinn an Lebenszufriedenheit aufweisen, gefolgt von den Personen mit sehr großem finanziellem Spielraum. Analysiert man jedoch den Einkommenseffekt auf der Basis multipel imputierter Haushaltseinkommen, so kann man keine signifikanten Unterschiede zwischen den armen und nicht-armen Haushalten feststellen. Mit singularer Imputation wird bei den nordirischen Daten die Signifikanz der Gruppenunterschiede überschätzt und führt zu qualitativ anderen Schlußfolgerungen.

Dieser Befund wiederholt sich auch bei der Analyse der belgischen Daten. Mit singularer Imputation erhalten wir hochsignifikante Zufriedenheitsunterschiede zwischen den Personen in armen und nicht-armen Haushalten, wobei der Abstand zwischen armen und reichen Haushalten am geringsten ist, was als ein Hinweis auf Veenhovens Grenznutzenhypothese interpretiert werden kann. Analysiert man jedoch die multipel imputierten Daten und berücksichtigt bei der Berechnung des Signifikanzniveaus auch die imputationsbedingte Unsicherheit, so kann man keine signifikanten Unterschiede zwischen den Einkommensgruppen feststellen.

Insgesamt kann man sagen, daß diese vier Datensätze aufgrund eines hohen Ausfallanteils und eines hohen Prognosefehlers besonders stark von der Imputationsunsicherheit betroffen sind. Beim Signifikanztest kommt man in neun von 12 Paarvergleichen zu signifikanten Ergebnissen bei singulärer Imputation und nicht-signifikanten Ergebnissen bei multipler Imputation.

**Tabelle 4a: Signifikanzniveau (p) der Zufriedenheitseffekte ökonomischer Haushaltslagen in Belgien, Großbritannien, Irland und Nordirland, nach Imputationsverfahren, logistische Regressionsschätzung**

	Belgien		Großbritannien		Irland		Nordirland	
Ökonomische Lage des Haushalts	sing. Imp	mult. Imp	sing. Imp	mult. Imp	sing. Imp	mult. Imp	sing. Imp	mult. Imp
BESCHEIDEN	0,00	0,13	0,06	0,24	0,22	0,37	0,01	0,13
WOHLHABEND	0,00	0,08	0,00	0,04	0,04	0,26	0,00	0,13
REICH	0,05	0,10	0,00	0,00	0,01	0,09	0,01	0,10

**Tabelle 4b: Zufriedenheitseffekte ( $\exp(b) \cdot 100$ ) ökonomischer Haushaltslagen in Belgien, Großbritannien, Irland und Nordirland, nach Imputationsverfahren, logistische Regressionsschätzung**

	Belgien		Großbritannien		Irland		Nordirland	
Ökonomische Lage des Haushalts	sing. Imp	mult. Imp	sing. Imp	mult. Imp	sing. Imp	mult. Imp	sing. Imp	mult. Imp
BESCHEIDEN	154*	49	49	37	36	14	211*	109
WOHLHABEND	183*	97	178*	119*	84*	38	846*	165
REICH	116*	84	1040*	654*	163*	113	434*	394

\*  $p \leq 0,05$

Wie sehen nun die Folgen des Imputationsverfahrens bei jenen Daten aus, die in geringem Ausmaß von Ausfallrate und Prognoseunsicherheit betroffen sind? Im Falle Spaniens und Luxemburgs, die zwar eine relativ hohe Ausfallrate zwischen 32 und 39 Prozent fehlender Einkommenswerte, aber einen relativ geringen Prognosefehler aufweisen, führt die multiple Imputation zu keinen anderen Schlußfolgerungen hinsichtlich Signifikanz und Größenrelation der Haushaltsgruppenunterschiede (siehe Tabellen 5a und 5b).

**Tabelle 5a: Signifikanzniveaus (p) der Zufriedenheitseffekte ökonomischer Haushaltslagen in Spanien und Luxemburg, nach Imputationsverfahren, logistische Regressionsschätzung**

Ökonomische Lage des Haushalts	Spanien		Luxemburg	
	single Imp.	multip. Imp.	single Imp.	multip. Imp.
BESCHEIDEN	0,31	0,20	0,26	0,88
WOHLHABEND	0,02	0,04	0,34	0,46
REICH	0,59	0,55	0,65	0,70

**Tabelle 5b: Zufriedenheitseffekte (exp(b)\*100) ökonomischer Haushaltslagen in Spanien und Luxemburg, nach Imputationsverfahren, logistische Regressionsschätzung**

Ökonomische Lage des Haushalts	Spanien		Luxemburg	
	single Imp.	multip. Imp.	single Imp.	multip. Imp.
BESCHEIDEN	32	36	145	12
WOHLHABEND	106*	81*	122	115
REICH	22	26	78	56

\*  $p \leq 0.05$

Bei den Stichproben aus den Niederlanden, Griechenland und Italien, die zwar eine mittlere Ausfallrate zwischen 14 und 25 Prozent aufweisen, aber einen relativ hohen Prognosefehler zwischen 40 und 44, sind inkonsistente Schätzungen zu beobachten (siehe Tabellen 6a und 6b).

Für Niederlande erhalten wir mit multipel imputierten Daten signifikante Unterschiede zwischen den Haushaltsgruppen, mit singulär imputierten Daten jedoch nicht. Die Größenrelation der Parameter ist jedoch konsistent.

Im Fall der italienischen Daten ist ein gegenteiliges Ergebnis festzustellen. Bezüglich des Unterschiedes zwischen Personen in armen Haushalten und Personen in Haushalten mit relativ bescheidenem Einkommen würden wir mit singulär imputierten Daten einen Unterschied feststellen ( $p \leq .05$ ) bei multipel imputierten Daten jedoch nicht ( $p > .10$ ). Es fällt auf, daß diese Inkonsistenzen dann nicht auftreten, wenn die Gruppenunterschiede sehr groß sind.

Bei den griechischen Daten schließlich ist hinsichtlich der Signifikanz der Gruppenunterschiede keine Inkonsistenz festzustellen, wohl aber in der Größenrelation der Parameter. Bei singulär imputierten Daten würde man im Falle Griechenlands einen mit dem Haushaltseinkommen linear wachsenden Zugewinn an Lebenszufriedenheit feststellen. Mit multipel imputierten Daten müßte man jedoch festhalten, daß ab einem bestimmten Ni-

veau wachsender Wohlstand in Griechenland nicht zu einem weiteren Gewinn an Lebenszufriedenheit führt.

**Tabelle 6a: Signifikanzniveaus (p) der Zufriedenheitseffekte ökonomischer Haushaltslagen in Niederlande, Italien und Griechenland, nach Imputationsverfahren, logistische Regressionsschätzung**

	Niederlande		Italien		Griechenland	
Ökonomische Lage des Haushalts	singl. Imp.	mult. Imp.	sing. Imp.	mult. Imp.	sing. Imp.	mult. Imp.
BESCHEIDEN	0,16	0,03	0,04	0,11	0,01	0,03
WOHLHABEND	0,00	0,00	0,01	0,03	0,00	0,00
REICH	0,13	0,01	0,00	0,00	0,00	0,03

**Tabelle 6b: Zufriedenheitseffekte ( $\exp(b)*100$ ) ökonomischer Haushaltslagen in Niederlande, Italien und Griechenland, nach Imputationsverfahren, logistische Regressionsschätzung**

	Niederlande		Italien		Griechenland	
ökonomische Lage des Haushalts	sing. Imp.	mult. Imp.	sing. Imp.	mult. Imp.	sing. Imp.	mult. Imp.
BESCHEIDEN	54	102*	89*	49	63*	52*
WOHLHABEND	644*	521*	142*	84*	143*	119*
REICH	102	221*	514*	211*	256*	126*

\*  $p \leq 0,05$

Bei den Ländern schließlich, deren Daten sowohl mit geringerer Ausfallrate als auch mit geringerem Prognosefehler behaftet sind, können wir insgesamt keine Inkonsistenz der Schlußfolgerungen auf der Basis singulär imputierter Daten gegenüber multipel imputierter Daten feststellen. Dies gilt sowohl für die Signifikanz als auch für die Größenrelation der Gruppenunterschiede (siehe Tabellen 7a und 7b).

**Tabelle 7a: Signifikanzniveaus (p) der Zufriedenheitseffekte ökonomischer Haushaltslagen, nach Imputationsverfahren, logistische Regressionsschätzung**

	Portugal		Ostdeut		Westdeut		Norwegen		Dänemark		Frankreich	
Lage des Haushalts	s. I.	m. I.	s. I.	m. I.	s. I.	m. I.	s. I.	m. I.	s. I.	m. I.	s. I.	m. I.
BESCHIEDEN	0,00	0,03	0,01	0,02	0,54	0,27	0,22	0,25	0,03	0,10	0,70	0,57
WOHLHAB.	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01	0,01	0,23	0,41
REICH	0,01	0,01	0,80	0,65	0,01	0,00	0,10	0,10	0,57	0,74	0,35	0,34

**Tabelle 7b: Zufriedenheitseffekte (exp(b)\*100) ökonomischer Haushaltslagen, nach Imputationsverfahren, logistische Regressionsschätzung**

	Portugal		Ostdeut		Westdeut		Norwegen		Dänemark		Frankreich	
Lage des Haushalts	s. I.	m. I.	s. I.	m. I.	s. I.	m. I.	s. I.	m. I.	s. I.	m. I.	s. I.	m. I.
BESCHIEDEN	106*	134*	62*	58*	21	43	41	38	147*	109	12	18
WOHLHAB.	402*	368*	156*	142*	227*	198*	132*	129*	210*	218*	52	42
REICH	164*	159*	37	67	360*	442*	185*	189*	778	540	54	52

\* $p \leq 0.05$

#### 4. Zusammenfassung

Wir haben in der vorliegenden Analyse eine einfache Hypothese über den Einfluß des Haushaltseinkommens auf die Lebenszufriedenheit herangezogen, um zwei grundsätzlich unterschiedliche Strategien der Behandlung fehlender Einkommenswerte in ihren Folgen für die Überprüfung der Hypothese zu vergleichen.

Zunächst konnten wir feststellen, daß die Einkommensmessung beim Eurobarometer 1992 von einem hohen Anteil fehlender Werte betroffen ist. Die Einkommensvariable ist lückenhaft und mit einem je nach nationaler Stichprobe unterschiedlich hohen, insgesamt beträchtlichen Ausfall behaftet. Diese Unsicherheit der Einkommensmessung könnte man nur dann bei inhaltlichen Analysen ignorieren, wenn davon ausgegangen werden könnte, daß die durch Verweigerung und Weiß-Nicht-Reaktionen entstehenden Lücken im Daten-

satz rein zufällig entstanden sind. In diesem Falle könnten die mit Datenlücken behafteten Fälle von der Analyse ausgeklammert werden, ohne daß Folgen für die Interpretation der Ergebnisse befürchtet werden müßten. Diese Annahme ist aber in bezug auf die Einkommensangaben aufgrund theoretischer Überlegungen und empirischer Ergebnisse nicht zutreffend. Ein Ausklammern der Fälle mit fehlender Einkommensangabe läßt vor diesem Hintergrund befürchten, daß die inhaltlichen Ergebnisse verzerrt sind.

Wir konnten feststellen, daß in den unterschiedlichen nationalen Stichproben, zusätzlich zu den variierenden Ausfallraten auch unterschiedlich hohe Varianzen der Residuen in der Einkommensgleichung auftreten. D.h. daß sich in den einzelnen nationalen Stichproben des Eurobarometers schlechte Datenqualität und unzureichendes Modell in besonders gravierender Weise kombinieren. Man könnte erwarten, daß eine Kompensation des Ausfalls durch eine Imputation in diesen Fällen besonders irreführende Resultate hervorbringt.

Beim Vergleich der Parameterschätzungen, die auf der Basis von singulärer und multipler Imputation erzielt wurden, konnten wir feststellen, daß die Art der Imputation zu diametral entgegengesetzten substantiellen Schlußfolgerungen bezüglich des Brutto-Einkommenseffektes auf die Lebenszufriedenheit führen kann. Diese Inkonsistenz war besonders deutlich in Stichproben feststellbar, bei denen sich hohe Ausfallrate und hoher Fehler der Einkommensprognose kombinierten. Hier gab es in neun von 12 Paarvergleichen unterschiedliche Inferenzschätzungen. Bei Stichproben mit hohem Prognosefehler des Einkommensmodells und mittlerer Ausfallrate war in drei von 12 Paarvergleichen ein Unterschied zwischen singulärer und multipler Imputation festzustellen.

In den Stichproben aber, in denen die Varianz der Residuen moderat ist, unterscheiden sich die Parameterschätzung bei singulärer und multipler Datenimputation nicht.

## Anmerkungen

\*) Unser herzlicher Dank gilt Nanny Wermuth und Peter Schmidt, die mit konstruktiver Kritik dieser Arbeit zum Abschluß verholfen haben.

1) Zur Interpretation der b-Koeffizienten beachte man, daß  $\exp(b_i)$  der multiplikative Effekt ist, mit dem bei Anwachsen der zugehörigen Variablen  $X_i$  um eine Einheit die Chancen für das Ereignis, überwiegend zufrieden zu sein, verändert wird. Dabei werden die Chancen in Einheiten des Verhältnisses der Wahrscheinlichkeit des Ereignisses zu der komplementären Gegenwahrscheinlichkeit („odds“) gemessen. Die Veränderung wird häufig als prozentuale Abweichung im Vergleich zu einer Referenzgruppe ausgedrückt.



2) In der vorliegenden Analyse beschränken wir die substantielle Betrachtung auf den einfachen Effekt der Haushaltseinkommenslage auf die Lebenszufriedenheit und sprechen deshalb von einem Bruttoeffekt des Einkommens auf die Lebenszufriedenheit.

3) Die Auswahl der Determinanten des Einkommensmodells orientiert sich an den Arbeiten der Sozialstrukturforschung (Hradil 1987) und Konsumforschung (Kroeber-Riel/Weinberg 1996), die auf Bildung, Berufsposition, Alter und Geschlecht und Haushaltsausstattung mit hochwertigen Gütern bzw. Immobilien als Korrelate des Haushaltseinkommens hinweisen.

## Literatur

Hagenaars, A.J.M./De Vos, C./Zaidi, M.A., 1995, Armutsstatistik Ende der 80er Jahre: Untersuchungen auf der Basis von Mikrodaten. Brüssel, Luxemburg.

Hradil, S., 1987: Sozialstrukturanalyse in einer fortgeschrittenen Gesellschaft. Von Klassen und Schichten zu Lagen und Milieus. Opladen: Westdeutscher Verlag.

Inglehart, R., 1989: Kultureller Umbruch. Wertewandel in der westlichen Welt. Frankfurt a. M./New York: Campus.

Inglehart, R./Rabier, J.-R., 1986: Aspirations Adapt to Situations - But Why Are the Belgians so Much Happier than the French? A Cross-Cultural Analysis of the Subjective Quality of Life. S. 1-56 in: Andrews, F. N. (Hrsg.), Research on the Quality of Life. Ann Arbor.

Kommission der Europäischen Gemeinschaften (Hrsg.), 1992: Eurobarometer. Die öffentliche Meinung in der Europäischen Gemeinschaft. Nr. 37. Brüssel.

Kroeber-Riel, W./Weinberg, P., 1996: Konsumentenverhalten, München: Vahlen Verlag.

Little, R. J./Rubin, D. B., 1987: Statistical Analysis with Missing Data. New York: Wiley.

Little, R. J./Schenker, N., 1995: Missing Data. S. 39-75 in: Arminger, G./Clogg, C. C./Sobel, M. E. (Hrsg.), Handbook of Statistical Modeling in the Social and Behavioral Sciences. New York: Wiley.

Papastefanou, G./Fleck, M., 1995: Die Behandlung von fehlenden Einkommenswerten in Analysen des ALLBUS. Interner Arbeitsbericht, Abteilung 'Einkommen und Verbrauch', ZUMA, Mannheim.

Reinecke, J./Schmidt, P., 1996: Model Specification and Missing Value Treatment in Panel Data: Testing the Theory of Planned Behavior in a Three-wave Panel Study. In: Engel, U./Reinecke, J. (Eds.), Analysis of Change. Advanced Techniques in Panel Data Analysis. De Gruyter, Berlin.

Rubin, D. B., 1987: Multiple Imputation for Nonresponse in Surveys. New York: Wiley.

Rubin, D. B., 1995: Multiple Imputation After 18+ Years. Manuskript.

Rubin, D. B./Schenker, N., 1991: Multiple Imputation in Health-Care Databases: An Overview and Some Applications. *Statistics in Medicine* 10: 585-598.

Veenhoven, R., 1990: Inequality in Happiness. Inequality in Countries Compared Between Countries. Paper presented at the 12th Work Congress of Sociology in Madrid, July 1990. Working Group Social Indicators and Quality of Life. Session 7, Social Trends and Inequality, Rotterdam: Erasmus University.

Veenhoven, R., 1993: Happiness in Nations: Subjective Appreciation of Life in 56 Nations 1946 - 1992. Rotterdam, Erasmus University Rotterdam, RISBO-(Studies in Socio-Cultural Transformation; No. 2).

Veenhoven, R., 1994: The Cross-National Pattern of Happiness. Test of Predictions Implied in Three Theories of Happiness. Rotterdam: Erasmus University.

Yates, F., 1963, The analysis of replicated experiments when the field results are incomplete, *Empire Journal of Experimental Agriculture*, 1, 129-142.